

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:16:38 ; Search time 91.75 seconds
(without alignments)
15.135 Million cell updates/sec

Title: US-09-331-631a-3_COPY_74_116
Perfect score: 250
Sequence: 1 NQDDPQTCQCQCRRCRQGE.....RQDQYCRCKEICEEEEE 43

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	40.0	588	1	VCLB_GOSHI
2	90	36.0	605	1	P09801 gossypium h
3	75	30.0	47	1	P09799 gossypium h
4	75	30.0	154	1	AGRP_LURCY
5	63.5	25.4	33	1	Y551_CAEEL
6	63.5	25.4	284	1	MBP1_MAIZE
7	61.5	24.6	50	1	HSP1_MOUSE
8	61.5	24.6	50	1	HSP1_MOUSE
9	61	24.4	919	1	ANDR_HUMAN
10	60	24.0	907	1	ANDR_HUMAN
11	60	24.0	911	1	ANDR_PANTR
12	59	23.6	2318	1	NTC3_MOUSE
13	58	23.2	445	1	NRH3_MOUSE
14	58	23.2	447	1	NRH3_MOUSE
15	58	23.2	447	1	NRH3_MOUSE
16	57.5	23.0	217	1	KAPC_DICDI
17	57.5	23.0	284	1	KSI_HYDAT
18	57.5	23.0	284	1	TPM1_RAT
19	57.5	23.0	284	1	TPM2_HUMAN
20	57.5	23.0	284	1	TPM2_HUMAN
21	57.5	23.0	284	1	TPM2_HUMAN
22	57	22.8	217	1	EGFR_HUMAN
23	57	22.8	217	1	YOM6_CAEEL
24	57	22.8	1166	1	CAS2_SHEEP
25	56.5	22.6	284	1	XMRK_XIPMA
26	56.5	22.6	284	1	TPM1_CHICK
27	56.5	22.6	1407	1	TPMB_CHICK
28	56	22.4	210	1	TRHY_RABIT
29	56	22.4	251	1	YOEI_ECOLI
30	56	22.4	302	1	GDBX_WHEAT
31	56	22.4	327	1	GDBX_WHEAT
32	56	22.4	1210	1	EGFR_MOUSE
33	55.5	22.2	639	1	GLCX_SOYBN

34	55.5	22.2	968	1	CC28_SCHPO	Q10752 schizosacch
35	54.5	21.8	338	1	FSA_HUMAN	P19883 homo sapien
36	54	21.6	223	1	CAS2_CAPIT	P33049 capra hircu
37	54	21.6	258	1	ZSS_RICCO	P01089 ricinus com
38	54	21.6	285	1	INVO_CANFA	P18174 canis faml
39	54	21.6	577	1	MTG8_MOUSE	Q61909 mus musculu
40	54	21.6	604	1	MTG8_HUMAN	Q06455 homo sapien
41	54	21.6	626	1	AH12_ARAHY	P43238 arachis hyp
42	54	21.6	758	1	YM38_YEAST	Q03825 saccharomyc
43	54	21.6	1339	1	ERM3_RAT	O62799 rattus norv
44	54	21.6	1403	1	PRO_DROME	P29617 drosophila
45	53.5	21.4	49	1	HSP1_PIG	P04101 sus scrofa

ALIGNMENTS

RESULT 1	VCLB_GOSHI	STANDARD;	PRT;	588 AA.
AC	P09801;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	VICILIN C72 PRECURSOR (ALPHA-GLOBULIN B).			
OS	Gossypium hirsutum (Upland cotton).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Malvales; Malvaceae; Gossypium.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Chlan C.A., Pyle J.B., Legocki A.B., Dure L. III;			
RT	"Developmental biochemistry of cottonseed embryogenesis and germination. XVII. cDNA and amino acid sequences of the members of the storage protein families."			
RT	Plant Mol. Biol. 7:475-489(1986).			
CC	- FUNCTION: SEED STORAGE PROTEIN.			
CC	- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN BODIES.			
CC	- SIMILARITY: TO OTHER 75 SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONVICTILIN, CONGLICININ, ETC.).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	-----			
CC	EMBL: M16891; AAA33071.1; -.			
DR	PIR: A30838; FMCNAB.			
DR	HSSP: P50477; ICAX.			
DR	INTERPRO: IPR001113; -.			
DR	PFAM: PF00546; Seedstore-7s; 1.			
KW	Seed storage protein; Signal.			
FT	SIGNAL			
FT	CHAIN			
FT	SEQUENCE			

Query Match 40.0%; Score 100; DB 1; Length 588;
Best Local Similarity 45.0%; Pred. No. 0.00074;
Matches 18; Conservative 10; Mismatches 10; Indels 2; Gaps 1;

QY 3 DDQPTDCCQCRRCRQESGPRQDYCRCKEICEEEEE 42
DB 81 EDPRRYECCQCRRCRQEE--RQRPCCQCRCKRFEQEQ 118

RESULT 2
VCLB_GOSHI
ID VCLB_GOSHI STANDARD; PRT; 605 AA.

AC P09799;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE VICILIN GC72-A PRECURSOR (ALPHA-GLOBULIN A).
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Malvales; Malvaceae; Gossypium.
RN [1]
RP SEQUENCE FROM N.A.
RA Chian C.A., Borroto K., Kamalay J.A., Dure L. III;
RT "Developmental biochemistry of cottonseed embryogenesis and
RT germination. XIX. Sequences and genomic organization of the alpha
RT globulin (vicilin) genes of cottonseed.";
RL Plant Mol. Biol. 9:533-546(1987).
CC -1- FUNCTION: SEED STORAGE PROTEIN.
CC -1- SUBCELLULAR LOCATION: CYTOSOL; MEMBRANE-BOUND VACUOLAR PROTEIN
CC BODIES.
CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,
CC CONVICILIN, CONGLYCININ, ETC.).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M19378; AAA33069.1; -;
DR PIR; S06398; S06398.
DR HSSP; P50477; 1CAX.
DR INTERPRO; IPR001113; -;
DR PFAM; PF00546; Seedstore_7s; 1.
KW Seed storage protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 605 VICILIN GC72-A.
SQ SEQUENCE 605 AA; 71049 MW; C9DB9371C976953B CRC64;

Query Match 36.0%; Score 90; DB 1; Length 605;
Best Local Similarity 35.9%; Pred. No. 0.0085;
Matches 14; Conservative 10; Mismatches 15; Indels 0; Gaps 0;
OY 2 QDDPOTDCOCRCRQOESGPRQOQYCORCKEICEE 40
DB 115 QQQPDKRKECCQRCQWQDQRPKRKQCVKREGRQYQED 153

RESULT 3
AGRP_LUFCY
ID AGRP_LUFCY STANDARD: PRT: 47 AA.
AC P56568;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 6.5 KDA ARGININE/GLUTAMATE-RICH POLYPEPTIDE (6.5K-AGRP).
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Cucurbitales; Cucurbitaceae; Luffa.
RN [1]
RP SEQUENCE.
RA Kimura M., Park S.S., Sakai R., Yamasaki N., Funatsu G.;
RT "Primary structure of 6.5k-arginine/glutamate-rich polypeptide from
RT the seeds of sponge gourd (Luffa cylindrica).";
RL Biosci. Biotechnol. Biochem. 61:984-988(1997).
CC -1- FUNCTION: STORAGE PROTEIN WHICH PROVIDES NITROGEN AND CARBON
CC RESERVES DURING GERMINATION AND SEEDLING GROWTH.
CC -1- MASS SPECTROMETRY: MW=5693.39; METHOD=MALDI.

CC -1- SIMILARITY: SOME, TO 7S SEED STORAGE PROTEINS.
KW Seed storage protein.
FT DISULFID 12 33
FT DISULFID 16 29
SQ SEQUENCE 47 AA; 5698 MW; 588BDECB2273AC05 CRC64;

Query Match 30.0%; Score 75; DB 1; Length 47;
Best Local Similarity 37.1%; Pred. No. 0.041;
Matches 13; Conservative 10; Mismatches 8; Indels 4; Gaps 1;

OY 5 POTDCOCRCRCRQOESGPRQOQYCORCKEICEE 39
DB 5 PRTEYACRVRQCVAEHGVET---QRRCOQVCEK 35

RESULT 4
Y551_CAEEL
ID Y551_CAEEL STANDARD: PRT: 154 AA.
AC Q09662;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HYPOTHEICAL 15.6 KDA PROTEIN ZK673.1 IN CHROMOSOME II PRECURSOR.
GN ZK673.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL NZ;
RA McMurray A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL NZ;
RA Jones S.J.M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO C.ELEGANS E04D5.4 AND SOME, TO C.ELEGANS
CC C03G6.13.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Z48585; CAA88486.1; -;
DR HSSP; P01551; 1ACX.
DR WORMPEP; ZK673.1; CE16742.
KW Hypothetical protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 154 HYPOTHEICAL PROTEIN ZK673.1.
SQ SEQUENCE 154 AA; 15637 MW; 323A02C3A2E782AB CRC64;

Query Match 30.0%; Score 75; DB 1; Length 154;
Best Local Similarity 41.9%; Pred. No. 0.11;
Matches 13; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 3 DDPOTDCOCRCRQOESGPRQOQYCORCKEICEE 33
DB 69 DDPNTDCTQYSLCSNAKYTPLLQDFCKTC 99

RESULT 5
MBP1_MAIZE
ID MBP1_MAIZE STANDARD: PRT: 33 AA.
AC P28794;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)

```

DT 01-OCT-1996 (Rel. 34, last annotation update)
DE ANTIMICROBIAL PEPTIDE MBP-1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
CC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN 111
RE SEQUENCE.
RC STRAIN=CV, B73; TISSUE=SEED;
RX MEDLINE=92406801; PubMed=1527010;
RA Duvick J.P., Rood T., Rao A.G., Marshak D.R.;
RT "Purification and characterization of a novel antimicrobial peptide
from maize (Zea mays L.) kernels."
RL J. Biol. Chem. 267:118814-118820(1992).
CC -1- FUNCTION: INHIBITOR OF BOTH BACTERIAL AND FUNGAL GROWTH IN VITRO.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE EMBRYO PORTION OF THE
CC KERNEL.
DR PIR: A41822; A41822.
DR MAZEDB: 69182; -.
KW Fungicide; Antibiotic.
SQ SEQUENCE 33 AA; 4131 MW; B148F1B90E823599 CRC64;

Query Match 25.4%; Score 63.5; DB 1; Length 33;
Best Local Similarity 44.0%; Pred. No. 0.49;
Matches 11; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

OY 11 QCORRC-RQOESGPRQOQYCORRCK 34
Db 6 ECRROCLRRHGGPWETQECRRCR 30

RESULT 6
TPM1_MOUSE STANDARD; PRT; 284 AA.
AC P46901;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TROPOMYOSIN, FIBROBLAST ISOFORM 1 (TM-1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN 111
RE SEQUENCE FROM N.A.
RC STRAIN=C3H;
RX MEDLINE=92129366; PubMed=1733968;
RA Wang Y.C., Rubenstein P.A.;
RT "Choice of 3' cleavage/polyadenylation site in beta-tropomyosin RNA
processing is differentiation-dependent in mouse BC3H1 muscle
cells."
RL J. Biol. Chem. 267:2728-2736(1992).
CC -1- FUNCTION: THE FUNCTION OF TROPOMYOSIN IN SMOOTH MUSCLE AND NON-
CC MUSCLE CELLS IS NOT CLEAR.
CC -1- ALTERNATIVE PRODUCTS: SKELETAL MUSCLE BETA AND FIBROBLAST 1
CC ISOFORM ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: THE MOLECULE IS IN A COILED COIL STRUCTURE. THE SEQUENCE
CC EXHIBITS A PROMINENT SEVEN-RESIDUES PERIODICITY.
CC -1- SIMILARITY: BELONGS TO THE TROPOMYOSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X58381; CAA41271.1; -.
DR PIR: S23256; S23256.
DR SWISS-2DPAGE: P46901; MOUSE.
DR INTERPRO: IPR000533; -.
DR PRAM: PR00261; Tropomyosin. 1.
DR PRINTS: PR00194; TROPOMYOSIN.

```

```

DR PROSITE; P500326; TROPOMYOSIN; 1.
KW Coiled coil; Repeat; Multigene family; Alternative splicing.
SQ SEQUENCE 284 AA; 32944 MW; A8C0E5780DC6F6AF CRC64;

Query Match 25.4%; Score 63.5; DB 1; Length 284;
Best Local Similarity 34.1%; Pred. No. 2.8;
Matches 14; Conservative 10; Mismatches 14; Indels 3; Gaps 1;

OY 3 DDPOTDQOCRCRQOESGPRQOQYCORRCKEICEEERY 43
Db 23 DEADADKKADKCKOLE---EEQALQKKLKTDEVEERY 60

RESULT 7
HSP1_MOUSE STANDARD; PRT; 50 AA.
AC P02319;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SPERM PROTEIN P1 (CYSTEINE-RICH PROTEIN).
GN PRM1 OR PRM-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN 111
RE SEQUENCE FROM N.A.
RX MEDLINE=85199803; PubMed=2986684;
RA Kiene K.C., Distel R.J., Hecht N.B.;
RT "Nucleotide sequence of a cDNA clone encoding mouse protamine 1."
RL Biochemistry 24:719-722(1985).
RN 121
RE SEQUENCE FROM N.A.
RX MEDLINE=88193085; PubMed=3358932;
RA Johnson P.A., Peschon J.D., Yelick P.C., Palmiter R.D., Hecht N.B.;
RT "Sequence homologues in the mouse protamine 1 and 2 genes."
RL Biochim. Biophys. Acta 950:45-53(1988).
RN 131
RE SEQUENCE FROM N.A.
RX MEDLINE=87260978; PubMed=3037541;
RA Peschon J.D., Behringer R.R., Brinster R.L., Palmiter R.D.;
RT "Spermatid-specific expression of protamine 1 in transgenic mice."
RL Proc. Natl. Acad. Sci. U.S.A. 84:5316-5319(1987).
RN 141
RE SEQUENCE FROM N.A.
RX MEDLINE=88181903; PubMed=3445973;
RA Hecht N.B.;
RT "Gene expression during spermatogenesis."
RL Ann. N.Y. Acad. Sci. 513:90-101(1987).
RN 151
RE SEQUENCE FROM N.A.
RC STRAIN=C129;
RA Schlueter G., Engel W.;
RT Submitted (Jul-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PROTEIN SUBSTITUTES FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE
CC DNA-HELIX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02926; AAA39980.1; -.
DR EMBL: X07625; CAA30472.1; -.

```

```

DR EMBL: X14003; CAA32169.1; -
DR EMBL: M27500; AAA39985.1; -
DR EMBL: 247352; CAA87410.1; -
DR PIR: A02660; HSMSS1.
DR PIR: S03820; S03820.
DR PIR: A28331; A28331.
DR MGI: 97765; PRM1.
DR INTERPRO: IPR000221; -.
DR PFAM: PF00260; protamine_P1.1.
DR PROSITE: PS00048; protamine_P1.1.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KM Testis; DNA condensation; Nuclear protein.
FT INIT_MET 0 0
FT DISULFID 5 5 INTERCHAIN (WITH CYS-21) (BY SIMILARITY).
FT DISULFID 6 14 BY SIMILARITY.
FT DISULFID 21 21 INTERCHAIN (WITH CYS-5) (BY SIMILARITY).
FT DISULFID 36 36 INTERCHAIN (WITH CYS-36) (BY SIMILARITY).
FT DISULFID 37 47 BY SIMILARITY.
SQ SEQUENCE 50 AA: 6827 MW: B61CAF3D1BBF4978 CRC64:

Query Match 24.6%; Score 61.5; DB 1; Length 50;
Best Local Similarity 31.4%; Pred. No. 1.1;
Matches 11: Conservative 9; Mismatches 10; Indels 5; Gaps 1;

OY 9 COQCQRCHQDSGPRQOQYCORCKEICEEEY 43
   1:::||||: ||::|:|:| 1
DB 14 CRRRRRCRR-----RRRCRRRRRRRCRRRSY 43

RESULT 8
HSP1_RAT
ID HSP1_RAT STANDARD; PRT; 50 AA.
AC P10118;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE SPERM PROTAMINE P1 (CYSTEINE-RICH PROTAMINE).
GN PRM1 OR PRM-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89335257; PubMed=2757789;
RA Klemm U., Lee C.H., Burfield P., Hake S., Engel W.;
RT "Nucleotide sequence of a cDNA encoding rat protamine and the haploid
RL expression of the gene during rat spermatogenesis.";
RL Biol. Chem. Hoppe-Seyler 370:293-301(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96341725; PubMed=8720108;
RA Schueter G., Celik A., Obata R., Schlicker M., Hofferbert S.,
RA Schlung A., Adham I.M., Engel W.;
RT "Sequence analysis of the conserved protamine gene cluster shows that
RT it contains a fourth expressed gene.";
RL Mol. Reprod. Dev. 43:1-6(1996).
RN [3]
RP SEQUENCE.
RX MEDLINE=89207111; PubMed=3072011;
RA Ammer H., Henschen A.;
RT "Rat sperm protamine. Isolation and sequence analysis.";
RL Biol. Chem. Hoppe-Seyler 369:1301-1306(1988).
RN [4]
RP SEQUENCE OF 1-15 AND 44-50.
RX MEDLINE=76184789; PubMed=1268226;
RA Kistler W.S., Keim P.S., Hennikson R.L.;
RT "Partial structural analysis of the basic chromosomal protein of rat
RT spermatozoa.";
RL Biochim. Biophys. Acta 427:752-757(1976).
CC -I- FUNCTION: PROTAMINES SUBSTITUTES FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.

```

```

CC -I- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE
CC DNA-HELIX (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: NUCLEAR.
CC -I- TISSUE SPECIFICITY: TESTIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 246939; CAA87061.1; -
DR PIR: A27128; A27128.
DR PIR: S03997; S03997.
DR INTERPRO: IPR000221; -.
DR PFAM: PF00260; protamine_P1.1.
DR PROSITE: PS00048; protamine_P1.1.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KM Testis; DNA condensation; Nuclear protein.
FT INIT_MET 0 0
FT DISULFID 5 5 INTERCHAIN (WITH CYS-21) (BY SIMILARITY).
FT DISULFID 6 14 BY SIMILARITY.
FT DISULFID 21 21 INTERCHAIN (WITH CYS-5) (BY SIMILARITY).
FT DISULFID 36 36 INTERCHAIN (WITH CYS-36) (BY SIMILARITY).
FT DISULFID 37 47 BY SIMILARITY.
SQ SEQUENCE 50 AA: 6889 MW: B60B9F3D1BB6D978 CRC64:

Query Match 24.6%; Score 61.5; DB 1; Length 50;
Best Local Similarity 31.4%; Pred. No. 1.1;
Matches 11: Conservative 9; Mismatches 10; Indels 5; Gaps 1;

OY 9 COQCQRCHQDSGPRQOQYCORCKEICEEEY 43
   1:::||||: ||::|:|:| 1
DB 14 CRRRRRCRR-----RRRCRRRRRRRCRRRSY 43

RESULT 9
ANDR_HUMAN
ID ANDR_HUMAN STANDARD; PRT; 919 AA.
AC P10275;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANDROGEN RECEPTOR (DIHYDROTESTOSTERONE RECEPTOR).
GN AR OR NR3C4 OR DHTR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89112208; PubMed=3216866;
RA Lubahn D.B., Joseph D.R., Sar M., Tan J., Higgs H.N., Larson R.E.,
RA French F.S., Wilson E.M.;
RT "The human androgen receptor: complementary deoxyribonucleic acid
RT cloning, sequence analysis and gene expression in prostate.";
RL Mol. Endocrinol. 2:1265-1275(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9008302; PubMed=2594783;
RA Lubahn D.B., Brown T.R., Simental J.A., Higgs H.N., Migeon C.J.,
RA Wilson E.M., French F.S.;
RT "Sequence of the intron/exon junctions of the coding region of the
RT human androgen receptor gene and identification of a point mutation
RT in a family with complete androgen insensitivity.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:9534-9538(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90258935; PubMed=2342476;
RA Govindan M.V.;
RT "Specific region in hormone binding domain is essential for hormone

```

RT binding and trans-activation by human androgen receptor.";
RL Mol. Endocrinol. 4:417-427(1990).
[4]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RX MEDLINE=89017168; PubMed=2174628;
RA Chang C., Kokontis J., Liao S.;
RT "Structural analysis of complementary DNA and amino acid sequences of
human and rat androgen receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:7211-7215(1988).
[5]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RX MEDLINE=89098909; PubMed=2911578;
RA Tilley W.D., Marcelli M., Wilson J.D., McPhaul M.J.;
RT "Characterization and expression of a cDNA encoding the human androgen
receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:327-331(1989).
[6]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RX MEDLINE=91155943; PubMed=2293020;
RA Marcelli M., Tilley W.D., Wilson C.M., Griffin J.E., Wilson J.D.,
McPhaul M.J.;
RT "Definition of the human androgen receptor gene structure permits the
identification of mutations that cause androgen resistance: premature
termination of the receptor protein at amino acid residue 588 causes
complete androgen resistance.";
RL Mol. Endocrinol. 4:1105-1116(1990).
[7]
RP SEQUENCE OF 189-919 FROM N.A.
RX MEDLINE=88178111; PubMed=3353726;
RA Chang C., Kokontis J., Liao S.;
RT "Molecular cloning of human and rat complementary DNA encoding
androgen receptors.";
RL Science 240:324-326(1988).
[8]
RP SEQUENCE OF 468-919 FROM N.A.
RX MEDLINE=88204047; PubMed=337788;
RA Trapman J., Klaassen P., Kuiper G.G.J.M., van der Korput J.A.G.M.,
Faber P.W., van Rooij H.C.J., Geurts van Kessel A., Voorhorst M.M.,
Mulder E., Brinkmann A.O.;
RT "Cloning, structure and expression of a cDNA encoding the human
androgen receptor.";
RL Biochem. Biophys. Res. Commun. 153:241-248(1988).
[9]
RP POLYMORPHISM OF POLY-GLN REGION.
RX MEDLINE=92220629; PubMed=1561105;
RA Stedens H.F., Oostera B.A., Brinkmann A.O., Trapman J.;
RT "Trinucleotide repeat polymorphism in the androgen receptor gene
(AR).";
RL Nucleic Acids Res. 20:1427-1427(1992).
[10]
RP POLYMORPHISM OF POLY-GLY REGION.
RT TISSUE=BL00D;
RA Lu J., Danielsen M.;
RN Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
[11]
RP VARIANTS SBMA IN POLY-GLN REGION.
RX MEDLINE=91287825; PubMed=2062380;
RA Spada A.R., Wilson E.M., Lubahn D.B., Harding A.E., Fischback K.H.;
RT "Androgen receptor gene mutations in X-linked spinal and bulbar
muscular atrophy.";
RL Nature 352:77-79(1991).
[12]
RP REVIEW ON VARIANTS.
RX MEDLINE=95023089; PubMed=7937057;
RA Patterson M.N., Hughes L.A., Gottlieb B., Pinsky L.;
RT "The androgen receptor gene mutations database.";
RL Nucleic Acids Res. 22:3560-3562(1994).
[13]
RP REVIEW ON VARIANTS.
RX MEDLINE=97169385; PubMed=90165528;
RA Gottlieb B., Trifiro M., Lumbroso R., Vasilion D.M., Pinsky L.;
RT "The androgen receptor gene mutations database.";
RL Nucleic Acids Res. 25:158-162(1997).
[14]
RP VARIANT LNCAP ALA-877.
RX MEDLINE=91083633; PubMed=2260966;
RA Veldscholte J., Ris-Stalpers C., Kuiper G.G., Jenster G.,
Barreroets C., Klaassen E., van Rooij H.C.J., Trapman J.,
Brinkmann A.O., Mulder E.;
RT "A mutation in the ligand binding domain of the androgen receptor of
human LNCap cells affects steroid binding characteristics and
response to anti-androgens.";
RL Biochem. Biophys. Res. Commun. 173:534-540(1990).
[15]
RP VARIANT CAIS MET-866.
RX MEDLINE=91186983; PubMed=2082179;
RA Brown T.R., Lubahn D.B., Wilson E.M., French F.S., Migeon C.J.,
Corten J.L.;
RT "Functional characterization of naturally occurring mutant androgen
receptors from subjects with complete androgen insensitivity.";
RL Mol. Endocrinol. 4:1759-1772(1990).
[16]
RP VARIANT CYS-774.
RX MEDLINE=91310758; PubMed=1856263;
RA Marcelli M., Tilley W.D., Zoppi S., Griffin J.E., Wilson J.D.,
McPhaul M.J.;
RT "Androgen resistance associated with a mutation of the androgen
receptor at amino acid 772 (Arg-->Cys) results from a combination of
decreased messenger ribonucleic acid levels and impairment of
receptor function.";
RL J. Clin. Endocrinol. Metab. 73:318-325(1991).
[17]
RP VARIANT CAIS ASN-695 AND HIS-695, AND SEQUENCE OF 629-723 FROM N.A.
RX MEDLINE=92131007; PubMed=1775137;
RA Ris-Stalpers C., Trifiro M.A., Kuiper G.G., Jenster G., Romalo G.,
Sai T., van Rooij H.C., Kautman M., Rosenfield R.L., Liao S.;
RT "Substitution of aspartic acid-686 by histidine or asparagine in the
human androgen receptor leads to a functionally inactive protein with
altered hormone-binding characteristics.";
RL Mol. Endocrinol. 5:1562-1569(1991).
[18]
RP VARIANTS CAIS AND PAIS.
RX MEDLINE=93338440; PubMed=1307250;
RA Batch J.A., Williams D.M., Davies H.R., Brown B.D., Evans B.A.J.,
Hughes I.A., Patterson M.N.;
RT "Androgen receptor gene mutations identified by SSCP in fourteen
subjects with androgen insensitivity syndrome.";
RL Hum. Mol. Genet. 1:497-503(1992).
[19]
RP VARIANT CAIS VAL-787.
RX MEDLINE=92235226; PubMed=1569163;
RA Nakao R., Hajl M., Yanase T., Ogo A., Takayanagi R., Katsube T.,
Fukumaki Y., Nawata H.;
RT "A single amino acid substitution (Met-786-->Val) in the steroid-
binding domain of human androgen receptor leads to complete androgen
insensitivity syndrome.";
RL J. Clin. Endocrinol. Metab. 74:1152-1157(1992).
[20]
RP VARIANT LNCAP ALA-877.
RX MEDLINE=92222955; PubMed=1562539;
RA Veldscholte J., Barreroets C.A., Ris-Stalpers C., Kuiper G.G.,
Jenster G., Trapman J., Brinkmann A.O., Mulder E.;
RT "The androgen receptor in LNCap cells contains a mutation in the
ligand binding domain which affects steroid binding characteristics
and response to antiandrogens.";
RL J. Steroid Biochem. Mol. Biol. 41:665-669(1992).
[21]
RP VARIANT MET-730.
RX MEDLINE=92335289; PubMed=1631125;
RA Newmark J.R., Hardy D.O., Tonb D.C., Carter B.S., Epstein J.L.,
Isaacs W.B., Brown T.R., Barrack E.R.;
RT "Androgen receptor gene mutations in human prostate cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6319-6323(1992).

RN [122]
 RP VARIANT CAIS VAL-754.
 RX MEDLINE-93372806; PubMed-8103398;
 RA Lobbaccaro J.-M., Lumbroso S., Ktari R., Dumas R., Sultan C.;
 RT "An exonic point mutation creates a Meili site in the androgen
 receptor gene of a family with complete androgen insensitivity
 syndrome.";
 RL Hum. Mol. Genet. 2:1041-1043(1993).
 RN [123]
 RP VARIANT CAIS ARG-807.
 RX MEDLINE-94108430; PubMed-8281140;
 RA Adeyemo O., Kallio P.J., Palvimo J.J., Kontula K., Jaenme O.A.;
 RT "A single-base substitution in exon 6 of the androgen receptor gene
 causing complete androgen insensitivity: the mutated receptor fails
 to transactivate but binds to DNA in vitro.";
 RL Hum. Mol. Genet. 2:1809-1812(1993).
 RN [124]
 RP VARIANT PAIS VAL-743.
 RX MEDLINE-93315568; PubMed-8325932;
 RA Nakao R., Yanase T., Sakai Y., Hagi M., Nawata H.;
 RT "A single amino acid substitution (Gly743 -> Val) in the steroid-
 binding domain of the human androgen receptor leads to Reifenstein

Query Match 24.4%; Score 61; DB 1; Length 919;
 Best Local Similarity 48.3%; Pred. No. 13;
 Matches 14; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 2 ODDPQTDCCQCRRCRQESGPRQOYCQ 30
 DB 61 QQQQQQQQQQQQQQQQQQETSPRQOQQQQ 89

RESULT 10
 ANDR_CANFA STANDARD; PRT: 907 AA.

AC 09T790;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ANDROGEN RECEPTOR (DIHYDROTESTOSTERONE RECEPTOR).
 GN AR OR NR3C4.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lu B., Smock S.L., Castleberry T.A., Owen T.A.;
 RT "Molecular cloning and functional characterization of the canine
 androgen receptor";
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
 PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC -I- SUBCELLULAR LOCATION: NUCLEAR.
 CC -I- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NR3 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF197950; AAF18084.1; -
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KM Zinc-finger; Steroid-binding.
 FT DOMAIN 1 545 MODULATING (BY SIMILARITY).
 FT DNA_BIND 547 612 C4-TYPE ZINC FINGERS (TWO).
 FT

FT ZN_FING 547 567 C4-TYPE.
 FT ZN_FING 583 607 C4-TYPE.
 FT DOMAIN 678 707 LIGAND-BINDING.
 FT DOMAIN 55 54 POLY-GLN.
 FT DOMAIN 70 76 POLY-GLN.
 FT DOMAIN 131 134 POLY-GLN.
 FT DOMAIN 202 202 POLY-GLN.
 FT DOMAIN 329 332 POLY-SER.
 FT DOMAIN 375 384 POLY-PRO.
 FT DOMAIN 399 405 POLY-ALA.
 SQ SEQUENCE 907 AA; 98726 MW; C6619F78DD2338AF CRC64;

Query Match 24.0%; Score 60; DB 1; Length 907;
 Best Local Similarity 48.1%; Pred. No. 16;
 Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 5 PQTDCCQCRRCRQESGPRQOYCQ 31
 DB 50 PGANHQQQQQQQQQQQQQQQQQETSPRQOQQQQ 76

RESULT 11
 ANDR_PANTR STANDARD; PRT: 911 AA.

AC 09T775;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ANDROGEN RECEPTOR (DIHYDROTESTOSTERONE RECEPTOR).
 GN AR OR NR3C4.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98404153; PubMed-9732460;
 RA Choong C.S., Kempainen J.A., Wilson E.M.;
 RT "Evolution of the primate androgen receptor: a structural basis for
 disease.";
 RL J. Mol. Evol. 47:334-342(1998).
 CC -I- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
 PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC -I- SUBCELLULAR LOCATION: NUCLEAR.
 CC -I- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NR3 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U94177; AAC73048.1; -
 DR HSSP: P06536; IRGD.
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001103; -
 DR INTERPRO: IPR001628; -
 DR PFAM: PF00104; hormone_rec; 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRINTS: PR00521; ANDROGENR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KM Zinc-finger; Steroid-binding.
 FT DOMAIN 1 549 MODULATING (BY SIMILARITY).
 FT DNA_BIND 551 616 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 551 571 C4-TYPE.


```

FT DISULFID 263 272 BY SIMILARITY.
FT DISULFID 279 292 BY SIMILARITY.
FT DISULFID 286 301 BY SIMILARITY.
FT DISULFID 303 312 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 324 339 BY SIMILARITY.
FT DISULFID 341 350 BY SIMILARITY.
FT DISULFID 356 367 BY SIMILARITY.
FT DISULFID 361 378 BY SIMILARITY.
FT DISULFID 380 389 BY SIMILARITY.
FT DISULFID 396 409 BY SIMILARITY.
FT DISULFID 403 418 BY SIMILARITY.
FT DISULFID 420 429 BY SIMILARITY.
FT DISULFID 436 447 BY SIMILARITY.
FT DISULFID 441 456 BY SIMILARITY.
FT DISULFID 458 467 BY SIMILARITY.
FT DISULFID 474 485 BY SIMILARITY.
FT DISULFID 479 494 BY SIMILARITY.
FT DISULFID 496 505 BY SIMILARITY.
FT DISULFID 512 523 BY SIMILARITY.
FT DISULFID 517 532 BY SIMILARITY.
FT DISULFID 534 543 BY SIMILARITY.
FT DISULFID 550 560 BY SIMILARITY.
FT DISULFID 555 569 BY SIMILARITY.
FT DISULFID 571 580 BY SIMILARITY.
FT DISULFID 587 598 BY SIMILARITY.
FT DISULFID 592 607 BY SIMILARITY.
FT DISULFID 609 618 BY SIMILARITY.
FT DISULFID 625 635 BY SIMILARITY.
FT DISULFID 630 644 BY SIMILARITY.
FT DISULFID 646 655 BY SIMILARITY.
FT DISULFID 662 673 BY SIMILARITY.
FT DISULFID 667 682 BY SIMILARITY.
FT DISULFID 684 693 BY SIMILARITY.
FT DISULFID 700 710 BY SIMILARITY.
FT DISULFID 705 719 BY SIMILARITY.
FT DISULFID 721 730 BY SIMILARITY.
FT DISULFID 739 750 BY SIMILARITY.
FT DISULFID 744 759 BY SIMILARITY.
FT DISULFID 761 770 BY SIMILARITY.
FT DISULFID 776 787 BY SIMILARITY.
FT DISULFID 781 797 BY SIMILARITY.
FT DISULFID 799 808 BY SIMILARITY.
FT DISULFID 815 827 BY SIMILARITY.
FT DISULFID 821 836 BY SIMILARITY.
FT DISULFID 838 847 BY SIMILARITY.
FT DISULFID 854 865 BY SIMILARITY.
FT DISULFID 859 874 BY SIMILARITY.
FT DISULFID 876 885 BY SIMILARITY.
FT DISULFID 892 902 BY SIMILARITY.
FT DISULFID 897 911 BY SIMILARITY.
FT DISULFID 913 922 BY SIMILARITY.
FT DISULFID 929 940 BY SIMILARITY.
FT DISULFID 934 949 BY SIMILARITY.
FT DISULFID 951 960 BY SIMILARITY.
FT DISULFID 967 978 BY SIMILARITY.
FT DISULFID 972 987 BY SIMILARITY.
FT DISULFID 989 998 BY SIMILARITY.
FT DISULFID 1005 1016 BY SIMILARITY.
FT DISULFID 1010 1023 BY SIMILARITY.
FT DISULFID 1025 1034 BY SIMILARITY.
FT DISULFID 1041 1062 BY SIMILARITY.
FT DISULFID 1056 1071 BY SIMILARITY.
FT DISULFID 1073 1082 BY SIMILARITY.
FT DISULFID 1089 1100 BY SIMILARITY.
FT DISULFID 1094 1109 BY SIMILARITY.
FT DISULFID 1111 1120 BY SIMILARITY.
FT DISULFID 1127 1138 BY SIMILARITY.
FT DISULFID 1132 1147 BY SIMILARITY.
FT DISULFID 1149 1158 BY SIMILARITY.
FT DISULFID 1165 1183 BY SIMILARITY.
FT DISULFID 1177 1192 BY SIMILARITY.
FT DISULFID 1194 1203 BY SIMILARITY.

```

```

FT DISULFID 1210 1223 BY SIMILARITY.
FT DISULFID 1215 1233 BY SIMILARITY.
FT DISULFID 1235 1244 BY SIMILARITY.

Query Match
Best Local Similarity 23.6%; Score 59; DB 1; Length 2318;
Matches 19; Conservative 5; Mismatches 17; Indels 32; Gaps 4;

OY 1 NODDPTDQ--QCOR-----RC-----RQESGPRQOVC----- 29
Db 1421 NVDDPWRQCEALQCRFLFNNSRCDPACSSPACLYNPDICYSGRDRTCNPIYKICADHF 1480
OY 30 -ORCKEICEEBE 41
Db 1481 ADGRCDQCGNTEE 1493

RESULT 13.
NRH3_MOUSE STANDARD; PRT; 445 AA.
AC Q920Y9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE OXYSTEROLS RECEPTOR LXR-ALPHA (LIVER X RECEPTOR ALPHA) (NUCLEAR ORPHAN
DE RECEPTOR LXR-ALPHA).
GN NR1H3 OR LXR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Y.E., Horiuchi M., Dzau V.J.;
RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: ORPHAN RECEPTOR. INTERACTION OF LXR ALPHA WITH RXR
CC SHIFTS RXR FROM ITS ROLE AS A SILENT DNA-BINDING PARTNER TO AN
CC ACTIVE LIGAND-BINDING SUBUNIT IN MEDIATING RETINOID RESPONSES
CC THROUGH TARGET GENES DEFINED BY LXRES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF085745; AAD16050.1; -.
CC HSSP; P20393; 1A6Y.
CC INTERPRO: IPR000536; -.
CC INTERPRO: IPR001628; -.
CC PFM; PF00104; hormone_rec.1.
CC PFM; PF00105; zf-C4.1.
CC PRINTS; PR00047; STROIDFINGER.
CC PROSITE; PS00031; NUCLEAR_RECEPTOR.1.
CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger.
FT DNA_BIND 96 161 C4-TYPE ZINC FINGERS (TWO).
FT ZN_FING 96 116 C4-TYPE.
FT ZN_FING 132 156 C4-TYPE.
SQ SEQUENCE 445 AA; 50417 MW; 1A426F38D935731 CRC64;

Query Match
Best Local Similarity 23.2%; Score 58; DB 1; Length 445;
Matches 16; Conservative 9; Mismatches 5; Indels 10; Gaps 4;

OY 9 COOCO-RRCROESGPRQOVC-----ORCKEICEEBE 42
Db 148 COECLRLKCRQ--AGWREE--CVLSSEBQIRLKLKROEBE 183

```

```

RESULT 14
ID NRH3_HUMAN STANDARD; PRT; 447 AA.
AC 013133;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE OXYSTEROLS RECEPTOR LXR-ALPHA (LIVER X RECEPTOR ALPHA) (NUCLEAR ORPHAN
DE RECEPTOR LXR-ALPHA).
GN NRH3 OR LXR-ALPHA.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=95262897; PubMed=7744246;
RA Wally P.J., Umesono K., Ong E.S., Evans R.M., Heyman R.A.,
RA Mangelsdorf D.J.;
RT "LXR, a nuclear receptor that defines a distinct retinoid response
RT pathway.";
RL Genes Dev. 9:1033-1045(1995).
CC -!- FUNCTION: ORPHAN RECEPTOR. INTERACTION OF LXR ALPHA WITH RXR
CC SHIFTS RXR FROM ITS ROLE AS A SILENT DNA-BINDING PARTNER TO AN
CC ACTIVE LIGAND-BINDING SUBUNIT IN MEDIATING RETINOID RESPONSES
CC THROUGH TARGET GENES DEFINED BY LXRES.
CC -!- SUBUNIT: INTERACTS WITH RXR..
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- INDUCTION: BY 9-CIS RETINOIC ACID (9CRA).
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NRI SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: U22662; AAA5856.1; -.
CC DR HSSP: P20393; IABY.
CC DR MIM: 602423; -.
CC DR INTERPRO: IPR000536; -.
CC DR INTERPRO: IPR001628; -.
CC DR PFAM: PF00104; hormone_rec; 1.
CC DR PFAM: PF00105; zf-C4; 1.
CC DR PRINTS: PR00047; STROIDFINGER.
CC DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
CC KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger.
CC FT DNA_BIND 98 163 C4-TYPE ZINC FINGERS (TWO).
CC FT ZN_FING 98 118 C4-TYPE.
CC FT ZN_FING 134 158 C4-TYPE.
CC SQ SEQUENCE 447 AA; 50480 MW; A227B233DB0F8C96 CRC64;

Query Match 23.2%; Score 58; DB 1; Length 447;
Best Local Similarity 40.0%; Pred. No. 15;
Matches 16; Conservative 9; Mismatches 5; Indels 10; Gaps 4;

```

```

DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT (EC 2.7.1.37).
GN PKAC OR PK2 OR PK3.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Dictyostellida; Dictyostellium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91323730; PubMed=1864510;
RA Buerki E., Anjard C., Scholder J.-C., Reymond C.D.;
RT "Isolation of two genes encoding putative protein kinases regulated
RT during Dictyostelium discoideum development.";
RL Gene 102:57-65(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93385090; PubMed=8373760;
RA Anjard C., Etcheberry L., Pinard S., Veron M., Reymond C.D.;
RT "An unusual catalytic subunit for the CAMP-dependent protein kinase
RT of Dictyostelium discoideum.";
RL Biochemistry 32:9532-9538(1993).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=AX3;
RX MEDLINE=93066311; PubMed=1332055;
RA Mann S.K.O., Yonemoto W.M., Taylor S.S., Firtel R.A.;
RT "DpkK3, which plays essential roles during Dictyostelium development,
RT encodes the catalytic subunit of CAMP-dependent protein kinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10701-10705(1992).
CC -!- FUNCTION: ESSENTIAL FOR DIFFERENTIATION AND FRUIT MORPHOGENESIS.
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
CC -!- SUBUNIT: IN DICTYOSTELIUM THE Holoenzyme IS A DIMER COMPOSED OF
CC A REGULATORY (R) AND A CATALYTIC (C) SUBUNIT. IN THE PRESENCE OF
CC CAMP IT DISSOCIATES INTO THE ACTIVE C SUBUNIT AND AN R MONOMER.
CC -!- DEVELOPMENTAL STAGE: CAMP ACTIVITY IS LOW IN VEGETATIVELY GROWING
CC AMOEBAE, INCREASES DURING DEVELOPMENT OF AGGREGATION AND REACHES
CC A MAXIMUM AT CULMINATION.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMP SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: M38703; -. NOT_ANNOTATED_CDS.
CC DR PIR: JQ1150; JQ1150.
CC DR HSSP: P05132; ZCPK.
CC DR DICTYDB: DB02030; PKAC.
CC DR INTERPRO: IPR000719; -.
CC DR INTERPRO: IPR000961; -.
CC DR INTERPRO: IPR002290; -.
CC DR PFAM: PF00069; pkinase; 1.
CC DR PFAM: PF00433; pkinase_C; 1.
CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
CC DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC KW Transferase; Serine/threonine-protein kinase; ATP-binding; CAMP;
CC Phosphorylation.
CC FT DOMAIN 58 64 ASN-RICH.
CC FT DOMAIN 136 223 GLN-RICH.
CC FT DOMAIN 233 250 THR-RICH.
CC FT DOMAIN 336 590 PROTEIN KINASE.
CC FT NP_BIND 342 350 ATP (BY SIMILARITY).
CC FT BINDING 365 365 ATP (BY SIMILARITY).
CC FT ACT_SITE 459 459 BY SIMILARITY.
CC FT MOD_RES 490 490 PHOSPHORYLATION (BY SIMILARITY).
CC SQ SEQUENCE 648 AA; 74458 MW; D0F9B3A48C58D084 CRC64;

Query Match 23.2%; Score 58; DB 1; Length 648;

```

Best Local Similarity 29.3%; Pred. No. 20;
Matches 12; Conservative 13; Mismatches 16; Indels

```
Qy      2 QDDPQTDCGCCQRRCRQOESPPROOYCQRRCKEICEEEE 42
      |  | | | | | | | | | | | | | | | | | | | |
Db     145 QQQPQQQQPQQQQPQQQQQQQQPQQQQQPQQQLQNNQNNQQQQ 185
```

Search completed: March 1, 2001, 16:16:39
Job time: 419 sec